

CALCITONIN GENE RELATED PEPTIDE RECEPTOR

This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, the use of such polynucleotides and polypeptides, as well as the production of such polynucleotides and polypeptides. More particularly, the polypeptide of the present invention is a human 7-transmembrane receptor which has been identified as a human Calcitonin Gene Related Peptide receptor, sometimes hereinafter referred to as "CGRP". The invention also relates to inhibiting the action of such polypeptides.

It is well established that many medically significant biological processes are mediated by proteins participating in signal transduction pathways that involve G-proteins and/or second messengers, e.g., CAMP (Lefkowitz, Nature, 351:353-354 (1991)). Herein these proteins are referred to as proteins participating in pathways with G-proteins or PPG proteins. Some examples of these proteins include the G-Protein Couple Receptor (GPCR), such as those for adrenergic agents and dopamine (Kobilka, B.K., et al., PNAS, 84:46-50 (1987); Kobilka, B.K., et al., Science, 238:650-656 (1987);

1005436-013002

Bunzow, J.R., et al., *Nature*, 336:783-787 (1988)), G-proteins themselves, effector proteins, e.g., phospholipase C, adenylyl cyclase, and phosphodiesterase, and actuator proteins, e.g., protein kinase A and protein kinase C (Simon, M.I., et al., *Science*, 252:802-8 (1991)).

For example, in one form of signal transduction, the effect of hormone binding is activation of an enzyme, adenylyl cyclase, inside the cell. Enzyme activation by hormones is dependent on the presence of the nucleotide GTP, and GTP also influences hormone binding. A G-protein connects the hormone receptors to adenylyl cyclase. G-protein was shown to exchange GTP for bound GDP when activated by hormone receptors. The GTP-carrying form then binds to an activated adenylyl cyclase. Hydrolysis of GTP to GDP, catalyzed by the G-protein itself, returns the G-protein to its basal, inactive form. Thus, the G-protein serves a dual role, as an intermediate that relays the signal from receptor to effector, and as a clock that controls the duration of the signal.

The deduced amino acid sequence of the receptor of the present invention demonstrates the presence of seven hydrophobic regions characteristic of the G-protein coupled receptors. Also, it may be a member of a distinct family of G-protein coupled receptors which include the receptors for additional hormones that belong to the secretin/glucagon family of peptides that have also been shown to exhibit significant sequence identity to human calcitonin receptors. Members of this family of peptides also include the receptors for vasoactive intestinal peptide (VIP), (Aishihara et al., *Embo J.*, 10:1635-1641 (1991), growth hormone releasing hormone (GHRH) (Mayo, K.E. et al., *Mol. Endocrinol.*, 6:1734-1744 (1992)), and glucagon-like peptide 1 (Thornes, B., *Pnas, USA*, 85:8641-8645 (1992)). - The cross-reactivity among the other members of this ligand family includes gastric-inhibitory peptide (GIP), pituitary adenylyl cyclase

activating peptide (PACAP), peptide with histidine as N-terminus and isoleucine as C-terminus (PHI) and helodermin suggests that the receptors for these ligands are also likely to belong to this family.

Analysis of the structural features of this new family of receptors reveals several conserved structural motifs, including a seven (7) transmembrane motif and a long amino-terminal exocyttoplasmic region containing conserved cysteines and several N-linked glycosylation sites, as well as an amino-terminal hydrophobic stretch that could function as a leader sequence.

The calcitonin family of regulatory peptides comprises five known members: calcitonin (CT), 2 calcitonin-gene-related peptides (alpha and beta-CGRP), islet amyloid polypeptide (IAPP or amylin), Salmon calcitonin (sCT) and adrenomedulin. These peptides have been implicated in the pathophysiology of osteoporosis, hypertension and maturity onset diabetes, respectively. CGRP is a peptide with a plethora of different actions. As a neurotransmitter having an effect on the cardiovascular system and the gastrointestinal tract, it can produce effects within seconds. As a neuromodulator in the nervous system, it produces effects over a few minutes. In inflammation it acts as an autocoid factor, causing vasodilation for over a period of hours. As a trophic factor acting in the CNS and on skeletal muscle it produces changes which last for many days. CGRP could also stimulate more than one second messenger pathway.

The CGRP polypeptide of the present invention has been found to specifically bind to CGRP and has a 55% amino acid sequence identity to a human calcitonin receptor.

In accordance with one aspect of the present invention, there are provided novel CGRP polypeptides, as well as biologically active and diagnostically or therapeutically

useful fragments, analogs and derivatives thereof. The CGRP receptor of the present invention is of human origin.

In accordance with another aspect of the present invention, there are provided isolated nucleic acid molecules encoding CGRP receptor polypeptides, including mRNAs, DNAs, cDNAs, genomic DNA, as well as biologically active and diagnostically or therapeutically useful fragments, analogs and derivatives thereof.

In accordance with a further aspect of the present invention, there is provided a process for producing such polypeptides by recombinant techniques comprising culturing recombinant prokaryotic and/or eukaryotic host cells, containing a CGRP receptor nucleic acid sequence, under conditions promoting expression of said protein and subsequent recovery of said protein.

In accordance with yet a further aspect of the present invention, there are provided antibodies against such polypeptides.

In accordance with another embodiment, there is provided a process for using the CGRP receptor polypeptide to screen for receptor antagonists and/or agonists and/or receptor ligands.

In accordance with still another embodiment of the present invention there is provided a process of using such agonists for therapeutic purposes, for example, to treat hypercalcemia, osteoporosis, Paget's disease, hypertension, obesity, coronary artery disease, hypercalcemia of malignancy, to stimulate immune responses, angiogenesis, and inhibition of superoxide radical production.

In accordance with another aspect of the present invention there is provided a process of using such antagonists for treating pain transmission, arthritis, maturity onset diabetes, chronic inflammation, migraine, and certain cancers.

In accordance with another aspect of the present invention, there are provided nucleic acid probes comprising nucleic acid molecules of sufficient length to specifically hybridize to CGRP receptor sequences.

In accordance with another aspect of the present invention, there is provided a method of diagnosing a disease or a susceptibility to a disease relating to a mutation in the CGRP receptor nucleic acid sequences and the receptor encoded by such nucleic acid sequences.

In accordance with yet a further aspect of the present invention, there are provided processes for utilizing such polypeptides, or polynucleotides encoding such polypeptides, for *in vitro* purposes related to scientific research, synthesis of DNA and manufacture of DNA vectors.

These and other aspects of the present invention should be apparent to those skilled in the art from the teachings herein.

The following drawings are illustrative of embodiments of the invention and are not meant to limit the scope of the invention as encompassed by the claims.

Figure 1 shows the cDNA sequence and the corresponding deduced amino acid sequence of the CGRP receptor polypeptide of the present invention. The initial methionine amino acid is part of an 21 amino acid putative leader sequence that ends in a Threonine (Thr) residue. The standard one-letter abbreviations for amino acids are used. Two cryptic ATG codons (underlined) are present at amino acid positions -1 and -11 upstream from the putative authentic initiation codon.

Figure 2. A hydropathy plot of the deduced protein indicates the presence of seven hydrophobic (transmembrane) regions and a putative N-terminal signal sequence.

Figure 3. Circular-map of the vector used for expression of the CGRP receptor protein in mammalian cells. The coding region of the cDNA was cloned within the

KpnI/BamHI sites of the mammalian expression vector CDN forming the construct pCDNHCGRP.

Figure 4. cAMP response in 293 cells transiently transfected with the CGRP receptor expression construct, pCDNHCGRP. (A) Untransfected cells (control), no change in cAMP response to 0.3 μ M human CGRP was observed. (B) Transfected cells, cAMP responses ranged from 10 to 15 fold in two separate experiments. Values represent the mean plus/minus SEM. This result indicates that the recombinant receptor will functionally couple to cAMP in response to CGRP exposure.

Figure 5. cAMP response in individual 293 stable cell lines transformed with the CGRP receptor cDNA. Cells were exposed to 0.3 μ M of human CGRP. Basal values for control 293 cells were 29 to 130 pmol/well. Clone number 22 showed the greatest response to human CGRP treatment (approximately 100-fold stimulation).

Figure 6. Dose response of human CGRP-mediated cAMP accumulation in 293 cells (clone 22 cells from Figure 5) stably transformed with the CGRP receptor cDNA. Each data point represents the mean of triplicate determinations. The control untransfected 293 cells showed two-fold stimulation at 1,000 nM concentration. The calculated EC_{50} value was 0.9 nM. This results shows that the recombinant receptor is functionally coupled to human CGRP.

Figure 7. Effect of the CGRP receptor antagonist, CGRP (8-37) on CGRP-mediated cAMP accumulation in 293 cells stably transformed with CGRP receptor cDNA. The cells in a 6-well plate were incubated with increasing concentrations of human CGRP in the absence of [□] or presence [■] of 100 nM human CGRP (8-37) for 10 mins. at 37°C. Cyclic AMP was assayed by RIA. Data are presented as means of triplicate determinations of two separate experiments. Inclusions of human CGRP (8-37) (100 nM) shifted the CGRP concentration response curve to the right in a competitive manner. Human

CGRP (8-37) by itself had no effect on cAMP accumulations (data not shown). the calculated pA_2 value was 7.8. This result confirms that the recombinant receptor is specifically a CGRP receptor.

Figure 8. Saturation binding. (A) Binding of [125 I] human CGRP to the membranes from 293 cells (clone 22) stably transformed with CGRP receptor cDNA. Equilibrium binding was measured at 25 °C for 60 min. Specific binding (■) was taken as difference between binding in the absence (□) or in the presence of 0.1 μ M unlabeled human CGRP (◆). Data are from one experiment representative of three independent experiments done in duplicate. (B) Scatchard plot of [125 I] CGRP binding. The observed K_d = 18.6 pM and B_{max} = 89 fmol/mg protein. The results indicate high affinity, low density binding.

Figure 9. Competition binding profiles. Competition curves for representative CGRP analogs against [125 I]CGRP binding to membranes prepared from 293 cells (clone 22 cells) stably transformed with the CGRP receptor cDNA. Each curve is representative of two to three independent experiments, each carried out in duplicate. In the competition studies the rank order of potency was human CGRP (■) > human CGRP (8-37) (□) > human adrenomedullin (ADM) (●) >>> salmon calcitonin (sCT) (○), human calcitonin (hCT) (△), human amylin (*), porcine vasoactive intestinal peptide (VIP) (⊗) and salmon calcitonin (8-32) (sCT (8-32)) (▲). These results demonstrate that the CGRP receptor has a pharmacology profile similar to the native CGRP receptor.

Figure 10. Amino acid sequence comparison of CGRP receptor polypeptide (top line) to the human calcitonin receptor polypeptide (bottom line) illustrating 72 % similarity and 55 % identity at the amino acid level.

Figure 11. Amino acid sequence comparison of CGRP receptor polypeptide (top line) to the rat calcitonin-like

receptor (bottom line) illustrating 91 identity at the amino acid level.

In accordance with an aspect of the present invention, there is provided an isolated nucleic acid (polynucleotide) which encodes for the mature polypeptide having the deduced amino acid sequence of Figure 1 (SEQ ID No. 2) or for the mature polypeptide encoded by the cDNA of the clone deposited as ATCC Deposit No. 75824 on June 24, 1994.

A polynucleotide encoding a polypeptide of the present invention may be found in lung, heart and kidney. The polynucleotide of this invention was discovered in a cDNA library derived from human synovium. It is structurally related to the G protein-coupled receptor family. It contains an open reading frame encoding a protein of 461 amino acid residues of which approximately the first 21 amino acids residues are the putative leader sequence such that the mature protein comprises 440 amino acids. The protein exhibits the highest degree of homology to a rat calcitonin-like receptor with 88 % identity and 93 % similarity over the entire amino acid sequence.

The polynucleotide of the present invention may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. The coding sequence which encodes the mature polypeptide may be identical to the coding sequence shown in Figure 1 (SEQ ID No. 1) or that of the deposited clone or may be a different coding sequence which coding sequence, as a result of the redundancy or degeneracy of the genetic code, encodes the same mature polypeptide as the DNA of Figure 1 (SEQ ID No. 1) or the deposited cDNA.

The polynucleotide which encodes for the mature polypeptide of Figure 1 (SEQ ID No. 2) or for the mature polypeptide encoded by the deposited cDNA may include: only

the coding sequence for the mature polypeptide; the coding sequence for the mature polypeptide and additional coding sequence such as a leader sequence; the coding sequence for the mature polypeptide (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the mature polypeptide.

Thus, the term "polynucleotide encoding a polypeptide" encompasses a polynucleotide which includes only coding sequence for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequence.

The present invention further relates to variants of the hereinabove described polynucleotides which encode for fragments, analogs and derivatives of the polypeptide having the deduced amino acid sequence of Figure 1 (SEQ ID No. 2) or the polypeptide encoded by the cDNA of the deposited clone. The variant of the polynucleotide may be a naturally occurring allelic variant of the polynucleotide or a non-naturally occurring variant of the polynucleotide.

Thus, the present invention includes polynucleotides encoding the same mature polypeptide as shown in Figure 1 (SEQ ID No. 2) or the same mature polypeptide encoded by the cDNA of the deposited clone as well as variants, either natural splicing variants or variants derived from recombinant DNA techniques, of such polynucleotides which variants encode for a fragment, derivative or analog of the polypeptide of Figure 1 (SEQ ID No. 2) or the polypeptide encoded by the cDNA of the deposited clone. Such nucleotide variants include deletion variants, substitution variants and addition or insertion variants.

As hereinabove indicated, the polynucleotide may have a coding sequence which is a naturally occurring allelic variant of the coding sequence shown in Figure 1 (SEQ ID No. 1) or of the coding sequence of the deposited clone. As known in the art, an allelic variant is an alternate form of

10054186 012202

a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded polypeptide.

The present invention also includes polynucleotides, wherein the coding sequence for the mature polypeptide may be fused in the same reading frame to a polynucleotide sequence which aids in expression and secretion of a polypeptide from a host cell, for example, a leader sequence which functions as a secretory sequence for controlling transport of a polypeptide, or a portion of a polypeptide, from the cell. The polynucleotides may also encode for a proprotein which is the mature protein plus additional 5' amino acid residues. A mature protein having a prosequence is a proprotein and is an inactive form of the protein. Once the prosequence is cleaved an active mature protein remains. A soluble form of the CGRP receptor may be the mature polypeptide less the putative leader sequence, or another portion of the mature polypeptide, that is not membrane bound.

The polynucleotides of the present invention may also have the coding sequence fused in frame to a marker sequence which allows for purification of the polypeptide of the present invention. The marker sequence may be a hexahistidine tag supplied by a pQE-9 vector to provide for purification of the mature polypeptide fused to the marker in the case of a bacterial host, or, for example, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson, I., et al., Cell, 37:767 (1984)).

The term "gene" means the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

10054136-012300
20210-9875007

Fragments of the full length gene of the present invention may be used as a hybridization probe for a cDNA library to isolate the full length cDNA and to isolate other cDNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type preferably have at least 30 bases and may contain, for example, 50 or more bases. The probe may also be used to identify a cDNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete gene including regulatory and promotor regions, exons, and introns. An example of a screen comprises isolating the coding region of the gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The present invention further relates to polynucleotides which hybridize to the hereinabove-described sequences if there is at least 70%, preferably at least 90%, and more preferably at least 95% identity between the sequences. The present invention particularly relates to polynucleotides which hybridize under stringent conditions to the hereinabove-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. The polynucleotides which hybridize to the hereinabove described polynucleotides in a preferred embodiment encode polypeptides which either retain substantially the same biological function or activity as the mature polypeptide encoded by the cDNAs of Figure 1 (SEQ ID NO:1) or the deposited cDNA(s).

Alternatively, the polynucleotide may have at least 20 bases, preferably 30 bases, and more preferably at least 50 bases which hybridize to a polynucleotide of the present

invention and which has an identity thereto, as hereinabove described, and which may or may not retain activity. For example, such polynucleotides may be employed as probes for the polynucleotide of SEQ ID NO:1, for example, for recovery of the polynucleotide or as a diagnostic probe or as a PCR primer.

Thus, the present invention is directed to polynucleotides having at least a 70% identity, preferably at least 90% and more preferably at least a 95% identity to a polynucleotide which encodes the polypeptide of SEQ ID NO:2 as well as fragments thereof, which fragments have at least 30 bases and preferably at least 50 bases and to polypeptides encoded by such polynucleotides.

The deposit(s) referred to herein will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for purposes of Patent Procedure. These deposits are provided merely as convenience to those of skill in the art and are not an admission that a deposit is required under 35 U.S.C. §112. The sequence of the polynucleotides contained in the deposited materials, as well as the amino acid sequence of the polypeptides encoded thereby, are incorporated herein by reference and are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

The present invention further relates to a CGRP receptor polypeptide which has the deduced amino acid sequence of Figure 1 (SEQ ID No. 2) or which has the amino acid sequence encoded by the deposited cDNA, as well as fragments, analogs and derivatives of such polypeptide.

The terms "fragment," "derivative" and "analog" when referring to the polypeptide of Figure 1 (SEQ ID No. 2) or that encoded by the deposited cDNA, means a polypeptide which either retains substantially the same biological function or

activity as such polypeptide, i.e. functions as a CGRP receptor, or retains the ability to bind the ligand or the receptor even though the polypeptide does not function as a CGRP receptor, for example, a soluble form of the receptor. An analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide, preferably a recombinant polypeptide.

The fragment, derivative or analog of the polypeptide of Figure 1 (SEQ ID No. 2) or that encoded by the deposited cDNA may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or

polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

The polypeptides of the present invention include the polypeptide of SEQ ID NO:2 (in particular the mature polypeptide) as well as polypeptides which have at least 70% similarity (preferably at least 70% identity) to the polypeptide of SEQ ID NO:2 and more preferably at least 90% similarity (more preferably at least 90% identity) to the polypeptide of SEQ ID NO:2 and still more preferably at least 95% similarity (still more preferably at least 90% identity) to the polypeptide of SEQ ID NO:2 and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 amino acids and more preferably at least 50 amino acids.

As known in the art "similarity" between two polypeptides is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide.

Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention. The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the CRT genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The polynucleotides of the present invention may be employed for producing polypeptides by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing a polypeptide. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the E. coli. lac or trp, the phage lambda P₁

105186 012203
promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in E. coli.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as E. coli, Streptomyces, Salmonella typhimurium; fungal cells, such as yeast; insect cells such as Drosophila S2 and Spodoptera Sf9; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided

by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and PCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, P_L and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of

appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples including the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), α -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation, initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein, or a portion thereof, into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include E. coli, Bacillus subtilis, Salmonella typhimurium and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell

lysing agents, such methods are well know to those skilled in the art.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa, 293 and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

The CGRP receptor polypeptides can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The polypeptides of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. ←

Polypeptides of the invention may also include an initial methionine amino acid residue.

The receptor of the present invention binds CGRP and has high amino acid sequence homology to calcitonin receptors. Accordingly, it may be employed in a process for screening for antagonists and/or agonists (both peptide and non-peptide in nature) for the receptor.

In general, such screening procedures involve providing appropriate cells which express the receptor on the surface thereof. In particular, a polynucleotide encoding the receptor of the present invention is employed to transfect cells to thereby express the CGRP receptor on the surface thereof. Such transfection may be accomplished by procedures as hereinabove described.

One such screening procedure involves the use of melanophores which are transfected to express the CGRP receptor of the present invention. Such a screening technique is described in PCT WO 92/01810 published February 6, 1992, which is herein incorporated by reference.

Thus, for example, such assay may be employed for screening for a receptor antagonist by contacting the melanophore cells which encode the CGRP receptor with both the receptor ligand and a compound to be screened. Inhibition of the signal generated by the ligand indicates that a compound is a potential antagonist for the receptor, i.e., inhibits activation of the receptor.

The screen may be employed for determining an agonist by contacting such cells with compounds to be screened and determining whether such compound generates a signal, i.e., activates the receptor to stimulate the accumulation of cAMP.

Other screening techniques include the use of cells which express the CGRP receptor (for example, transfected CHO cells) in a system which measures extracellular pH changes caused by receptor activation, for example, as described in Science, volume 246, pages 181-296 (October 1989), herein

incorporated by reference. For example, potential agonists or antagonists may be contacted with a cell which expresses the CGRP receptor and a second messenger response, e.g. signal transduction or pH changes, or making use of a reporter gene system, for example luciferase, may be measured to determine whether the potential agonist or antagonist is effective.

Another such screening technique involves introducing RNA encoding the CGRP receptor into *Xenopus* oocytes to transiently express the receptor. The receptor oocytes may then be contacted in the case of antagonist screening with the receptor ligand and a compound to be screened, followed by detection of inhibition of a calcium or cAMP signal, or in the case of an agonist, by detection of stimulation of a calcium or cAMP signal.

Another screening technique involves expressing the CGRP receptor in which the receptor is linked to a phospholipase C or D. As representative examples of such cells, there may be mentioned endothelial cells, smooth muscle cells, embryonic kidney cells, etc. The screening for an antagonist or agonist may be accomplished as hereinabove described by detecting activation of the receptor or inhibition of activation of the receptor from the phospholipase second signal.

Another method involves screening for CGRP receptor inhibitors by determining inhibition of binding of labeled ligand to cells or membranes which have the receptor on the surface thereof. Such a method involves transfecting a eukaryotic cell with DNA encoding the CGRP receptor such that the cell expresses the receptor on its surface and contacting the cell with a potential antagonist in the presence of a labeled form of a known ligand. The ligand can be labeled, e.g., by radioactivity. The amount of labeled ligand bound to the receptors is measured, e.g., by measuring radioactivity of the receptors. If the potential antagonist

binds to the receptor, as determined by a reduction of labeled ligand which binds to the receptors, the binding of labeled ligand to the receptor is inhibited.

Another method involves screening for CGRP inhibitors by determining inhibition of CGRP-mediated cAMP and/or adenylate cyclase accumulation. Such a method involves transfecting a eukaryotic cell with CGRP receptor to express the receptor on the cell surface. The cell is then exposed to potential antagonists in the presence of CGRP. The amount of cAMP accumulation is then measured. If the potential antagonist binds the receptor, and thus inhibits CGRP binding, the levels of CGRP-mediated cAMP, or adenylate cyclase, activity will be reduced.

The present invention also provides a method for determining whether a ligand not known to be capable of binding to a CGRP receptor can bind to such receptor which comprises contacting a mammalian cell which expresses a CGRP receptor with the ligand under conditions permitting binding of ligands to the CGRP receptor, detecting the presence of a ligand which binds to the receptor and thereby determining whether the ligand binds to the CGRP receptor. The systems hereinabove described for determining agonists and/or antagonists may also be employed for determining ligands which bind to the receptor.

In general, agonists for CGRP receptors are employed for therapeutic purposes, such as the treatment of Parkinson's disease, acute heart failure, hypotension, urinary retention, and osteoporosis.

Antagonists for CGRP receptors may be employed for a variety of therapeutic purposes. For example, such antagonists have been employed for treatment of hypertension, angina pectoris, myocardial infarction, ulcers, asthma, allergies, psychoses, depression, migraine, vomiting, and benign prostatic hypertrophy.

Examples of potential CGRP receptor antagonists are an antibody, or in some cases an oligonucleotide, which binds to the G-protein coupled receptor but does not elicit a second messenger response such that the activity of the G-protein coupled receptor is prevented.

Potential antagonists also include proteins which are closely related to the ligand of the CGRP receptor, i.e. a fragment of the ligand, which have lost biological function and when binding to the CGRP receptor, elicit no response.

A potential antagonist also includes an antisense construct prepared through the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes for the mature polypeptides of the present invention, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix -see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al, Science, 241:456 (1988); and Dervan et al., Science, 251: 1360 (1991)), thereby preventing transcription and the production of CGRP receptor. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the CGRP receptor (antisense - Okano, J. Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the CGRP receptor.

Another potential antagonist is a small molecule which binds to the CGRP receptor, making it inaccessible to ligands such that normal biological activity is prevented. Examples

of small molecules include but are not limited to small peptides or peptide-like molecules.

Potential antagonists also include a soluble form of a CGRP receptor, e.g. a fragment of the receptor, which binds to the ligand and prevents the ligand from interacting with membrane bound CGRP receptors.

The agonists identified by the screening procedures outlined above may be employed to enhance the neuromodulatory functions of CGRP to treat hypertension since CGRP is a potent vasodilator which increases coronary blood flow. The agonists may also be used to increase coronary blood flow in occluded coronary vessels. Similarly, CGRP has powerful excitatory effects on cardiac contractility and may be employed to stimulate heart contractility. Agonists may also be employed to treat gigantism by decreasing growth hormone release.

The agonists may also be employed to treat osteoporosis since CGRP inhibits osteoclast-mediated bone resorption and stimulates osteogenesis. In this same manner hypercalcemia may be treated. Similarly, the agonists may be employed to treat Paget's Disease.

The agonists may also be employed to stimulate angiogenesis and promote wound healing via the stimulatory effect of CGRP on endothelial cell proliferation.

The agonists may also be employed to treat obesity since CGRP controls feeding behavior by decreasing appetite and intestinal motility.

The agonists may also be employed to stimulate nerve regeneration since CGRP is a trophic agent in the CNS.

The agonists may also be employed to enhance the immune response through increasing vascular permeability.

Agonists may also be employed to inhibit superoxide production. Superoxide production is known in the art to cause cellular damage and lead possibly to diseases such as cancer.

The CGRP receptor antagonists may also be employed to inhibit CNS pain transmission, to treat chronic inflammation caused by long-lived vasodilation, arthritis, maturity onset diabetes, cardiovascular disorders and to treat migraine headaches.

The antagonists may also be employed to prevent carcinoid tumor of the lung, since an elevated level of CGRP has been found in the lung during this disease.

The antagonists and agonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as hereinafter described.

CGRP nucleic acid sequences and CGRP peptides may also be employed for *in vitro* purposes related to scientific research, synthesis of DNA and manufacture of DNA vectors, and for the production of diagnostics and therapeutics to treat human disease.

Fragments of the full length CGRP receptor gene may be used as a hybridization probe for a cDNA library to isolate the full length CGRP receptor gene and to isolate other genes which have a high sequence similarity to the gene or similar biological activity. Probes of this type generally have at least 20 bases. Preferably, however, the probes have at least 30 bases and generally do not exceed 1000 base pairs, although they may have a greater number of bases. The probe may also be used to identify a cDNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete CGRP receptor gene including regulatory and promotor regions, exons, and introns. An example of a screen comprises isolating the coding region of the CGRP receptor gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The CGRP receptor polypeptides and antagonists or agonists which are polypeptides, may be employed in accordance with the present invention by expression of such polypeptides *in vivo*, which is often referred to as "gene therapy."

Thus, for example, cells from a patient may be engineered with a polynucleotide (DNA or RNA) encoding a polypeptide *ex vivo*, with the engineered cells then being provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, cells may be engineered by procedures known in the art by use of a retroviral particle containing RNA encoding a polypeptide of the present invention.

Similarly, cells may be engineered *in vivo* for expression of a polypeptide *in vivo* by, for example, procedures known in the art. As known in the art, a producer cell for producing a retroviral particle containing RNA encoding the polypeptide of the present invention may be administered to a patient for engineering cells *in vivo* and expression of the polypeptide *in vivo*. These and other methods for administering a polypeptide of the present invention by such method should be apparent to those skilled in the art from the teachings of the present invention. For example, the expression vehicle for engineering cells may be other than a retrovirus, for example, an adenovirus which may be used to engineer cells *in vivo* after combination with a suitable delivery vehicle.

Retroviruses from which the retroviral plasmid vectors hereinabove mentioned may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

In one embodiment, the retroviral plasmid vector is derived from Moloney Murine Leukemia Virus.

The vector includes one or more promoters. Suitable promoters which may be employed include, but are not limited to, the retroviral LTR; the SV40 promoter; and the human cytomegalovirus (CMV) promoter described in Miller, et al., Biotechniques, Vol. 7, No. 9, 980-990 (1989), or any other promoter (e.g., cellular promoters such as eukaryotic cellular promoters including, but not limited to, the histone, pol III, and β -actin promoters). Other viral promoters which may be employed include, but are not limited to, adenovirus promoters, thymidine kinase (TK) promoters, and B19 parvovirus promoters. The selection of a suitable promoter will be apparent to those skilled in the art from the teachings contained herein.

The nucleic acid sequence encoding the polypeptide of the present invention is under the control of a suitable promoter. Suitable promoters which may be employed include, but are not limited to, adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; the ApoAI promoter; human globin promoters; viral thymidine kinase promoters, such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs (including the modified retroviral LTRs hereinabove described); the β -actin promoter; and human growth hormone promoters. The promoter also may be the native promoter which controls the gene encoding the polypeptide.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, ψ -2, ψ -AM, PA12, T19-14X, VT-19-17-H2, ψ CRE, ψ CRIP, GP+E-86, GP+envAm12, and DAN cell

lines as described in Miller, Human Gene Therapy, Vol. 1, pgs. 5-14 (1990), which is incorporated herein by reference in its entirety. The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and CaPO_4 precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line generates infectious retroviral vector particles which include the nucleic acid sequence(s) encoding the polypeptides. Such retroviral vector particles then may be employed, to transduce eukaryotic cells, either *in vitro* or *in vivo*. The transduced eukaryotic cells will express the nucleic acid sequence(s) encoding the polypeptide. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, as well as hematopoietic stem cells, hepatocytes, fibroblasts, myoblasts, keratinocytes, endothelial cells, and bronchial epithelial cells.

The CGRP receptor polypeptides and antagonists or agonists, may be employed in combination with a suitable pharmaceutical carrier. Such compositions comprise a therapeutically effective amount of the polypeptide or antagonist or agonist, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The formulation should suit the mode of administration.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological

products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the pharmaceutical compositions may be employed in conjunction with other therapeutic compounds.

The pharmaceutical compositions may be administered in a convenient manner such as by the topical, intravenous, intraperitoneal, intramuscular, intranasal or intradermal routes. The pharmaceutical compositions are administered in an amount which is effective for treating and/or prophylaxis of the specific indication. In general, the pharmaceutical compositions will be administered in an amount of at least about 10 $\mu\text{g/kg}$ body weight and in most cases they will be administered in an amount not in excess of about 8 mg/Kg body weight per day. In most cases, the dosage is from about 10 $\mu\text{g/kg}$ to about 1 mg/kg body weight daily, taking into account the routes of administration, symptoms, etc.

This invention further provides a method of screening drugs or identifying drugs which specifically interact with, and bind to, the human CGRP receptor polypeptide on the surface of a cell which comprises contacting a mammalian cell comprising an isolated DNA molecule encoding the CGRP receptor with a plurality of drugs, determining those drugs which bind to the mammalian cell, and thereby identifying drugs which specifically interact with and bind to a CGRP receptor polypeptide of the present invention.

This invention also provides a method of detecting the expression of the CGRP receptor on the surface of a cell by detecting the presence of mRNA encoding a CGRP receptor polypeptide which comprises obtaining total mRNA from the cell and contacting the mRNA so obtained with a nucleic acid probe comprising a nucleic acid molecule of at least 15 nucleotides capable of specifically hybridizing with a sequence included within the sequence of a nucleic acid molecule encoding the human CGRP receptor under hybridizing conditions, detecting the presence of mRNA hybridized to the

probe, and thereby detecting the expression of the CGRP receptor by the cell.

This invention is also related to the use of the CGRP receptor gene as part of a diagnostic assay for detecting diseases or susceptibility to diseases related to the presence of mutations in CGRP receptor nucleic acid sequences.

Individuals carrying mutations in the CGRP receptor gene may be detected at the DNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from a patient's cells, such as from blood, urine, saliva, tissue biopsy and autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR (Saiki et al., Nature, 324:163-166 (1986)) prior to analysis. RNA or cDNA may also be used for the same purpose. As an example, PCR primers complementary to the nucleic acid encoding the CGRP receptor can be used to identify and analyze CGRP receptor mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled CGRP receptor RNA or alternatively, radiolabeled CGRP receptor antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase A digestion or by differences in melting temperatures. Alternatively, point mutations may be detected by direct sequencing of the PCR products or sequencing of cloned PCR products.

Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution-gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of

different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (see, e.g., Myers et al., Science, 230:1242 (1985)).

Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (e.g., Cotton et al., PNAS, USA, 85:4397-4401 (1985)).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, (e.g., Restriction Fragment Length Polymorphisms (RFLP)) and Southern blotting of genomic DNA.

In addition to more conventional gel-electrophoresis and DNA sequencing, mutations can also be detected by *in situ* analysis.

The present invention also relates to a diagnostic assay for detecting altered levels of CGRP receptor protein in various tissues since an over-expression of the proteins compared to normal control tissue samples may detect the presence of a disease or susceptibility to a disease, for example, cancer. Assays used to detect levels of CGRP receptor protein in a sample derived from a host are well-known to those of skill in the art and include radioimmunoassays, competitive-binding assays, Western Blot analysis, ELISA assays and "sandwich" assay. An ELISA assay (Coligan, et al., Current Protocols in Immunology, 1(2), Chapter 6, (1991)) initially comprises preparing an antibody specific to the CGRP receptor antigen, preferably a monoclonal antibody. In addition a reporter antibody is prepared against the monoclonal antibody. To the reporter antibody is attached a detectable reagent such as radioactivity, fluorescence or, in this example, a horseradish peroxidase enzyme. A sample is removed from a host and incubated on a solid support, e.g. a polystyrene

10051186, 012202

dish, that binds the proteins in the sample. Any free protein binding sites on the dish are then covered by incubating with a non-specific protein like BSA. Next, the monoclonal antibody is incubated in the dish during which time the monoclonal antibodies attach to any CGRP receptor proteins attached to the polystyrene dish. All unbound monoclonal antibody is washed out with buffer. The reporter antibody linked to horseradish peroxidase is placed in the dish resulting in binding of the reporter antibody to any monoclonal antibody bound to CGRP receptor protein. Unattached reporter antibody is then washed out. Peroxidase substrates are then added to the dish and the amount of color developed in a given time period is a measurement of the amount of CGRP receptor protein present in a given volume of patient sample when compared against a standard curve.

A competition assay may be employed wherein antibodies specific to CGRP receptor polypeptide are attached to a solid support and labeled CGRP receptor and a sample derived from the host are passed over the solid support and the amount of label detected, for example by liquid scintillation chromatography, can be correlated to a quantity of CGRP receptor protein in the sample.

A "sandwich" assay is similar to an ELISA assay. In a "sandwich" assay CGRP receptor protein is passed over a solid support and binds to antibody attached to a solid support. A second antibody is then bound to the CGRP receptor protein. A third antibody which is labeled and specific to the second antibody is then passed over the solid support and binds to the second antibody and an amount can then be quantified.

The sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data

(repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3' untranslated regions is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 50 or 60 bases. For a review of this technique, see Verma et al., *Human Chromosomes: a Manual of Basic Techniques*, Pergamon Press, New York (1988).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, *Mendelian*

Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptides itself. In this manner, even a sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native

polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, 1975, Nature, 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole, et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

The present invention will be further described with reference to the following examples; however, it is to be understood that the present invention is not limited to such examples. All parts or amounts, unless otherwise specified, are by weight.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain

sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 μ g of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 μ l of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μ g of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37°C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

Size separation of the cleaved fragments is performed using 8 percent polyacrylamide gel described by Goeddel, D. et al., *Nucleic Acids Res.*, 8:4057 (1980).

"Oligonucleotides" refers to either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides have no 5' phosphate and thus will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide will ligate to a fragment that has not been dephosphorylated.

"Ligation" refers to the process of forming phosphodiester bonds between two double stranded nucleic acid fragments (Maniatis, T., et al., *Id.*, p. 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units to T4 DNA ligase ("ligase") per 0.5 μ g of approximately equimolar amounts of the DNA fragments to be ligated.

Unless otherwise stated, transformation was performed as described in the method of Graham, F. and Van der Eb, A., Virology, 52:456-457 (1973).

Example 1

Bacterial Expression and Purification of the CGRP receptor

The DNA sequence encoding CGRP receptor, ATCC # 75824, is initially amplified using PCR oligonucleotide primers corresponding to the 5' and 3' end sequences of the processed gene sequence (minus the signal peptide sequence) and the vector sequences 3' to the CGRP receptor gene. Additional nucleotides corresponding to the CGRP receptor coding sequence are added to the 5' and 3' sequences respectively. The 5' oligonucleotide primer has the sequence 5' GACTAAAGCTTAATGTTATACAGCATATTT 3' (SEQ ID No. 3) contains a HindIII restriction enzyme site followed by 18 nucleotides of the CGRP receptor coding sequence starting from the presumed terminal amino acid of the processed protein codon. The 3' sequence 5' GAACTTCTAGACCGTCAATTATATAAATTTTTC 3' (SEQ ID No. 4) contains complementary sequences to an XbaI site and is followed by 18 nucleotides of the CGRP receptor coding sequence. The restriction enzyme sites correspond to the restriction enzyme sites on the bacterial expression vector pQE-9 (Qiagen, Inc. Chatsworth, CA). pQE-9 encodes antibiotic resistance (Amp^r), a bacterial origin of replication (ori), an IPTG-regulatable promoter operator (P/O), a ribosome binding site (RBS), a 6-His tag and restriction enzyme sites. pQE-9 is then digested with HindIII and XbaI. The amplified sequences are ligated into pQE-9 and are inserted in frame with the sequence encoding for the histidine tag and the RBS. The ligation mixture is then used to transform E. coli strain M15/rep 4 (Qiagen, Inc.) by the procedure described in Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, Cold Spring Laboratory Press, (1989). M15/rep4 contains multiple copies

of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis. Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.⁶⁰⁰) of between 0.4 and 0.6. IPTG ("Isopropyl-B-D-thiogalacto pyranoside") is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression. Cells are grown an extra 3 to 4 hours. Cells are then harvested by centrifugation. The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl. After clarification, solubilized is purified from this solution by chromatography on a Nickel-Chelate column under conditions that allow for tight binding by proteins containing the 6-His tag (Hochuli, E. et al., J. Chromatography 411:177-184 (1984)). The CGRP receptor protein is eluted from the column in 6 molar guanidine HCl pH 5.0 and for the purpose of renaturation adjusted to 3 molar guanidine HCl, 100mM sodium phosphate, 10 mmolar glutathione (reduced) and 2 mmolar glutathione (oxidized). After incubation in this solution for 12 hours the protein is dialyzed to 10 mmolar sodium phosphate.

Example 2

Cloning and expression of CGRP receptor using the baculovirus expression system

The DNA sequence encoding the full length CGRP receptor protein, ATCC # 75824, is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene:

10051186"012200
202270"98TTSOT

The 5' primer has the sequence 5' GTCCGGATCCGCCACCATGTTATACAGCATATTT 3' (SEQ ID No. 5) and contains a BamHI restriction enzyme site (in bold) followed by 6 nucleotides resembling an efficient signal for the initiation of translation in eukaryotic cells (J. Mol. Biol. 1987, 196, 947-950, Kozak, M.), which is just upstream of the first 18 nucleotides of the CGRP receptor gene (the initiation codon for translation "ATG" is underlined).

The 3' primer has the sequence 5' GTCCGGATCCGCCACCATGTTATACAGCATATTT 3' (SEQ ID No. 6) and contains the cleavage site for the restriction endonuclease BamHI and 18 nucleotides complementary to the 3' non-translated sequence of the CGRP receptor gene. The amplified sequences are isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment is then digested with the endonuclease BamHI and then purified on a 1% agarose gel. This fragment is designated F2.

The vector pRG1 (modification of pVL941 vector, discussed below) is used for the expression of the CGRP receptor protein using the baculovirus expression system (for review see: Summers, M.D. and Smith, G.E. 1987, A manual of methods for baculovirus vectors and insect cell culture procedures, Texas Agricultural Experimental Station Bulletin No. 1555). This expression vector contains the strong polyhedrin promoter of the Autographa californica nuclear polyhedrosis virus (AcMNPV) followed by the recognition site for the restriction endonuclease BamHI. The polyadenylation site of the simian virus (SV)40 is used for efficient polyadenylation. For an easy selection of recombinant viruses the beta-galactosidase gene from E.coli is inserted in the same orientation as the polyhedrin promoter followed by the polyadenylation signal of the polyhedrin gene. The polyhedrin sequences are flanked at both sides by viral sequences for the cell-mediated homologous recombination of

10051136-012202

cotransfected wild-type viral DNA. Many other baculovirus vectors could be used in place of pRG1 such as pAc373, pVL941 and pAcIM1 (Luckow, V.A. and Summers, M.D., Virology, 170:31-39).

The plasmid is digested with the restriction enzyme BamHI and then dephosphorylated using calf intestinal phosphatase by procedures known in the art. The DNA is then isolated and purified on a 1% agarose gel as described above. This vector DNA is designated V2.

Fragment F2 and the dephosphorylated plasmid V2 are ligated with T4 DNA ligase. E.coli HB101 cells are then transformed and bacteria identified that contained the plasmid (pBacCGRP receptor) with the CGRP receptor gene using the enzyme BamHI. The sequence of the cloned fragment is confirmed by DNA sequencing.

5 μ g of the plasmid pBac-CGRP receptor are cotransfected with 1.0 μ g of a commercially available linearized baculovirus ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA.) using the lipofection method (Felgner et al. Proc. Natl. Acad. Sci. USA, 84:7413-7417 (1987)).

1 μ g of BaculoGold™ virus DNA and 5 μ g of the plasmid pBacCGRP receptor are mixed in a sterile well of a microtiter plate containing 50 μ l of serum free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards 10 μ l Lipofectin plus 90 μ l Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to the Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is rocked back and forth to mix the newly added solution. The plate is then incubated for 5 hours at 27°C. After 5 hours the transfection solution is removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. The plate is put back into an incubator and cultivation continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay performed similar as described by Summers and Smith (supra). As a modification an agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used which allows an easy isolation of blue stained plaques. (A detailed description of a "plaque assay" can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10).

Four days after the serial dilution, the viruses are added to the cells and blue stained plaques are picked with the tip of an Eppendorf pipette. The agar containing the recombinant viruses is then resuspended in an Eppendorf tube containing 200 μ l of Grace's medium. The agar is removed by a brief centrifugation and the supernatant containing the recombinant baculoviruses is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then stored at 4°C.

Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-CGRP receptor at a multiplicity of infection (MOI) of 2. Six hours later the medium is removed and replaced with SF900 II medium minus methionine and cysteine (Life Technologies Inc., Gaithersburg). 42 hours later 5 μ Ci of 35 S-methionine and 5 μ Ci 35 S cysteine (Amersham) are added. The cells are further incubated for 16 hours before they are harvested by centrifugation and the labelled proteins visualized by SDS-PAGE and autoradiography.

Example 3

Transient and Stable Expression in Cultured Cells

The DNA sequence encoding CGRP receptor polypeptide, ATCC # 75824, was re-configured for expression (with and without a C-terminal DET tag (KSIRIQRGPR)) in the mammalian expression vector pCDN. This involved the removal of the

entire 3' and 5' untranslated regions along with two cryptic methionine residues at putative amino acid positions -11 and -1 located upstream from the chosen methionine that could potentially negatively influence the expression. To increase the efficiency of expression, a flanking Kozak consensus sequence (CCACC) was added to the chosen start methionine residue. The resulting constructs were separately transfected into both COS and Human Kidney 293 cells by DEAE dextran sulfate and Lipofectin reagent protocols, respectively. For transient expression cells were assayed at 72 hours post-transfection for the accumulation of cAMP in response to ligand exposure. For one trial only the 293 cells demonstrated expression. For stable cell lines the transfected 293 cells were selected in the presence of 400 µg/ml of Geneticin (G418 Sulfate) and assayed for the accumulation of cAMP in response to CGRP.

Example 4

Expression via Gene Therapy

Fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin, is added. This is then incubated at 37°C for approximately one week. At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

PMV-7 (Kirschmeier, P.T. et al, DNA, 7:219-25 (1988) flanked by the long terminal repeats of the Moloney murine

10054186 012202

sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

The cDNA encoding a polypeptide of the present invention is amplified using PCR primers which correspond to the 5' and 3' end sequences respectively. The 5' primer containing an EcoRI site and the 3' primer further includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is used to transform bacteria HB101, which are then plated onto agar-containing kanamycin for the purpose of confirming that the vector had the gene of interest properly inserted.

The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells are transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it

is necessary to use a retroviral vector that has a selectable marker, such as neo or his.

The engineered fibroblasts are then injected into the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads. The fibroblasts now produce the protein product.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, within the scope of the appended claims, the invention may be practiced otherwise than as particularly described.

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: LI, ET AL.
- (ii) TITLE OF INVENTION: Calcitonin Gene Related Peptide Receptor
- (iii) NUMBER OF SEQUENCES: 6
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
- (B) STREET: 6 BECKER FARM ROAD
- (C) CITY: ROSELAND
- (D) STATE: NEW JERSEY
- (E) COUNTRY: USA
- (F) ZIP: 07068
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.5 INCH DISKETTE
- (B) COMPUTER: IBM PS/2
- (C) OPERATING SYSTEM: MS-DOS
- (D) SOFTWARE: WORD PERFECT 5.1
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE: Concurrently
- (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA
- (A) APPLICATION NUMBER: PCT/US94/09235
- (B) FILING DATE: August 16, 1994

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: FERRARO, GREGORY D.
(B) REGISTRATION NUMBER: 36,134
(C) REFERENCE/DOCKET NUMBER: 325800-376

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 201-994-1700
(B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 3034 BASE PAIRS
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CACGAGGGAA CAACCTCTCT CTCTSCAGCA GAGAGTGTCA CCTCCTGCTT TAGGACCATC 60
AAGCTCTGCT AACTGAATCT CATCCTAATT GCAGGATCAC ATTGCAAAGC TTTCACCTCTT 120
TCCCACCTTG CTTGTGGGTA AATCTCTTCT GCGGAATCTC AGAAAGTAAA GTTCCATCCTT 180
GAGAATATTT CACAAAGAAT TTCCTTAAGA GCTGGACTGG GTCTTGACCC CTGGAATTTA 240
AGAAATTCTT AAAGACAATG TCAAATATGA TCCAAGAGAA AATGTGATTT GAGTCTGGAG 300
ACAATTGTGC ATATCGTCTA ATAATAAAAA CCCATACTAG CCTATAGAAA ACAATATTTG 360
AATAATAAAA ACCCATACTA GCCTATAGAA AACAATATTT GAAAGATTGC TACCACTAAA 420
AAGAAAACCTA CTACAACTTG ACAAGACTGC TGCAAACCTC AATTGGTCAC CACAACCTGA 480
CAAGGTTGCT ATAAAACAAG ATTGCTACAA CTCTAGTTT ATGTTATACA GCATATTTCA 540
TTTGGGCTTA ATGATGGAGA AAAAGTGTAC CCTGTATTTT CTGTTTCTCT TGCCTTTTTT 600
TATGATTCTT GTTACAGCAG AATTAGAAGA GAGTCCCTGAG GACTCAATTC AGTTGGGAGT 660
TACTAGAAAT AAAATCATGA CAGCTCAATA TGAATGTTAC CAAAAGATTA TGCAAGACCC 720
CATTCAACAA GCAGAAGGCG TTTACTGCAA CAGAACCTGG GATGGATGGC TCTGCTGGAA 780
CGATGTTGCA GCAGGAAGT AATCAATGCA GCTCTGCCCT GATTACTTTC AGGACTTTGA 840
TCCATCAGAA AAAGTTACAA AGATCTGTGA CCAAGATGGA AACTGGTTTA GACATCCAGC 900
AAGCAACAGA ACATGGACAA ATTATACCCA GTGTAATGTT AACACCCACG AGAAAGTGAA 960
GACTGCACTA AATTTGTTTT ACCTGACCAT AATTGGACAC GGATTGTCTA TTGCATCACT 1020

GCTTATCTCG CTTGGCATAT TCTTTTATTT CAAGAGCCTA AGTTGCCAAA GGATTACCTT 1080
ACACAAAAAT CTGTTCTTCT CATTTGTTTG TAACTCTGTT GTAACAATCA TTCACCTCAC 1140
TGCAGTGGCC AACAACCAGG CCTTAGTAGC CACAAATCCT GTTAGTTGCA AAGTGTCCCA 1200
GTTCAATTCAT CTTTACCTGA TGGGCTGTAA TTACTTTTGG ATGCTCTGTG AAGGCATTTA 1260
CCTACACACA CTCATTGTGG TGGCCGTGTT TGCAGAGAAG CAACATTTAA TGTGGTATTA 1320
TTTTCTTGGC TGGGGATTTC CACTGATTCC TGCTTGTATA CATGCCATTG CTAGAAGCTT 1380
ATATTACAAT GACAATTGCT GGATCAGTTC TGATACCCAT CTCCTCTACA TTATCCATGG 1440
CCCAATTTGT GCTGCTTTAC TGGTGAATCT TTTTTTCTTG TTAAATATTG TACGCGTTCT 1500
CATCACCAAG TTAAAAGTTA CACACCAAGC GGAATCCAAT CTGTACATGA AAGCTGTGAG 1560
AGCTACTCTT ATCTTGGTGC CATTGCTTGG CATTGAATTT GTGCTGATTG CATGGCGACC 1620
TGAAGGAAAG ATTGCAGAGG AGGTATATGA CTACATCATG CACATCCTTA TGCACCTCCA 1680
GGGTCTTTTG GTCTCTACCA TTTTCTGCTT CTTTAATGGA GAGGTTCAAG CAATTCTGAG 1740
AAGAACTGG AATCAATACA AAATCCAATT TGGAAACAGC TTTTCCAACCT CAGAAGCTCT 1800
TCGTAGTGCG TCTTACACAG TGTCACAAT CAGTGATGGT CCAGGTTATA GTCATGACTG 1860
TCCTAGTGAA CACTTAAATG GAAAAAGCAT CCATGATATT GAAAATGTTT TCTTAAACC 1920
AGAAAATTTA TATAATTGAA AATAGAAGGA TGGTTGTCTC ACTGTTTGGT GCTTCTCCTA 1980
ACTCAAGGAC TTGGACCCAT GACTCTGTAG CCAGAAGACT TCAATATTAA ATGACTTTGG 2040
GGAATGTCAT AAAGAAGAGC CTTACATGA AATTAGTAGT GTGTTGATAA GAGTGTAACA 2100
TCCAGCTCTA TGTGGGAAAA AAGAAATCCT GGTGTTGTAAT GTTGTGAGT AAATACTCCC 2160
ACTATGCCTG ATGTGACGCT ACTAACCTGA CATCACCAAG TGTGGAATTG GAGAAAAGCA 2220
CAATCAACTT TTCTGAGCTG GTGTAAGCCA GTTCCAGCAC ACCATTGATG AATTCAAACA 2280
AATGGCTGTA AAACTAAACA TACATGTTGG GCATGATTCT ACCCTTATTG SCCCCAAGAG 2340
ACCTAGCTAA GGTCTATAAA CATGAAGGGA AAATTAGCTT TTAGTTTTAA AACTCTTTAT 2400
CCCATCTTGA TTGGGGCAGT TGACTTTTTT TTTTCCAG AGTGCCGTAG TCCTTTTTGT 2460
AACTACCCCTC TCAAATGGAC AATACCAGAA GTGAATTATC CCTGCTGGCT TTCTTTTCTC 2520
TATGAAAAGC AACTGAGTAC AATTGTTATG ATCTACTCAT TTGCTGACAC ATCAGTTATA 2580
TCTTGTGGCA TATCCATTGT GGAACTGGA TGAACAGGAT GTATAATATG CAATCTTACT 2640
TCTATATCAT TAGGAAAACA TCTTAGTTGA TGCTACAAAA CACCTTGTC AACTCTTCCT 2700
GTCTTACCAA ACAGTGGGAG GGAATTCCTA GCTGTAAATA TAAATTTTGC CCTTCCATT 2760
CTACTGTATA AACAAATTAG CAATCATTTT ATATAAGAA AATCAATGAA GGATTTCTTA 2820
TTTTCTTGA ATTTTGTA AAGAAATTGT GAAAATGAG CTGTAAATA CTCCATTAT 2880
TTATTTTATA GTCTCAAATC AAATACATAC AACCTATGTA ATTTTAAAG CAAATATATA 2940
ATGCAACAAT GTGTGTATGT TAATATCTGA TACTGTATCT GGGCTGATTT TTTAAATAAA 3000
ATAGAGTCTG GAATGCTAAA AAAAAAAAAA AAAA 3034

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 461 AMINO ACIDS

(B) TYPE: AMINO ACID

(C) STRANDEDNESS:

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Lys Lys Cys Thr Leu Tyr Phe Leu Val Leu Leu Pro Phe
-20 -15 -10
Phe Met Ile Leu Val Thr Ala Glu Leu Glu Glu Ser Pro Glu Asp
-5 1 5
Ser Ile Gln Leu Gly Val Thr Arg Asn Lys Ile Met Thr Ala Gln
10 15 20
Tyr Glu Cys Tyr Gln Lys Ile Met Gln Asp Pro Ile Gln Gln Ala
25 30 35
Glu Gly Val Tyr Cys Asn Arg Thr Trp Asp Gly Trp Leu Cys Trp
40 45 50
Asn Asp Val Ala Ala Gly Thr Glu Ser Met Gln Leu Cys Pro Asp
55 60 65
Tyr Phe Gln Asp Phe Asp Pro Ser Glu Lys Val Thr Lys Ile Cys
70 75 80
Asp Gln Asp Gly Asn Trp Phe Arg His Pro Ala Ser Asn Arg Thr
85 90 95
Trp Thr Asn Tyr Thr Gln Cys Asn Val Asn Thr His Glu Lys Val
100 105 110
Lys Thr Ala Leu Asn Leu Phe Tyr Leu Thr Ile Ile Gly His Gly
115 120 125
Leu Ser Ile Ala Ser Leu Leu-Ile-Ser Leu Gly Ile Phe Phe Tyr
130 135 140
Phe Lys Ser Leu Ser Cys Gln Arg Ile Thr Leu His Lys Asn Leu

145	150	155
Phe Phe Ser Phe Val Cys Asn Ser Val Val	Thr Ile Ile His Leu	
160	165	170
Thr Ala Val Ala Asn Asn Gln Ala Leu Val	Ala Thr Asn Pro Val	
175	180	185
Ser Cys Lys Val Ser Gln Phe Ile His Leu	Tyr Leu Met Gly Cys	
190	195	200
Asn Tyr Phe Trp Met Leu Cys Glu Gly Ile	Tyr Leu His Thr Leu	
205	210	215
Ile Val Val Ala Val Phe Ala Glu Lys Gln	His Leu Met Trp Tyr	
220	225	230
Tyr Phe Leu Gly Trp Gly Phe Pro Leu Ile	Pro Ala Cys Ile His	
235	240	245
Ala Ile Ala Arg Ser Leu Tyr Tyr Asn Asp	Asn Cys Trp Ile Ser	
250	255	260
Ser Asp Thr His Leu Leu Tyr Ile Ile His	Gly Pro Ile Cys Ala	
265	270	275
Ala Leu Leu Val Asn Leu Phe Phe Leu Leu	Asn Ile Val Arg Val	
280	285	290
Leu Ile Thr Lys Leu Lys Val Thr His Gln	Ala Glu Ser Asn Leu	
295	300	305
Tyr Met Lys Ala Val Arg Ala Thr Leu Ile	Leu Val Pro Leu Leu	
310	315	320
Gly Ile Glu Phe Val Leu Ile Pro Trp Arg	Pro Glu Gly Lys Ile	
325	330	335
Ala Glu Glu Val Tyr Asp Tyr Ile Met His	Ile Leu Met His Phe	
340	345	350
Gln Gly Leu Leu Val Ser Thr Ile Phe Cys	Phe Phe Asn Gly Glu	
355	360	365
Val Gln Ala Ile Leu Arg Arg Asn Trp Asn	Gln Tyr Lys Ile Gln	
370	375	380
Phe Gly Asn Ser Phe Ser Asn Ser Glu Ala	Leu Arg Ser Ala Ser	
385	390	395
Tyr Thr Val Ser Thr Ile Ser Asp Gly Pro	Gly Tyr Ser His Asp	
400	405	410

Cys Pro Ser Glu His Leu Asn Gly Lys Ser Ile His Asp Ile Glu
 415 420 425
 Asn Val Leu Leu Lys Pro Glu Asn Leu Tyr Asn
 430 435 440

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 30 BASE PAIRS
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GACTAAAGCT TAATGTTATA CAGCATATTT

30

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 33 BASE PAIRS
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GAACTTCTAG ACCGTCAATT ATATAAATTT TTC

33

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS
(A) LENGTH: 34 BASE PAIRS
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTCCGGATCC GCCACCATGT TATACAGCAT ATTT

34

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS
(A) LENGTH: 34 BASE PAIRS
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTCCGGATCC GCCACCATGT TATACAGCAT ATTT

34